Link Throughput of Multi-Channel Opportunistic Access with Limited Sensing

Keqin Liu, Qing Zhao

Department of Electrical and Computer Engineering

University of California, Davis, CA 95616

kqliu@ucdavis.edu, qzhao@ece.ucdavis.edu

Abstract

We aim to characterize the maximum link throughput of a multi-channel opportunistic communication system. The states of these channels evolve as independent and identically distributed Markov processes (the Gilbert-Elliot channel model). A user, with limited sensing and access capability, chooses one channel to sense and access in each slot and collects a reward determined by the state of the chosen channel. Such a problem arises in cognitive radio networks for spectrum overlay, opportunistic transmissions in fading environments, and resource-constrained jamming and anti-jamming. The objective of this report is to characterize the optimal performance of such systems. The problem can be generally formulated as obtaining the maximum expected long-term reward of a partially observable Markov decision process or a restless multi-armed bandit process, for which analytical characterizations are rare. Exploiting the structure and optimality of the myopic channel selection policy established recently, we obtain a closed-form expression of the maximum link throughput for two-channel systems and lower and upper bounds when there are more than two channels. These results allow us to study the rate at which the optimal performance of an opportunistic system increases with the number of channels and to obtain the limiting performance as the number of channels approaches to infinity.

Index Terms

Opportunistic access, cognitive radio, spectrum overlay, dynamic channel selection, myopic policy.

---

1. REPORT DATE **JUL 2007**	2. REPORT TYPE	3. DATES COVERED **00-00-2007 to 00-00-2007**

4. TITLE AND SUBTITLE **Link Throughput of Multi-Channel Opportunistic Access with Limited Sensing**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of California,Department of Electrical and Computer Engineering,Davis,CA,95616**	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**

13. SUPPLEMENTARY NOTES

14. ABSTRACT

We aim to characterize the maximum link throughput of a multi-channel opportunistic communication system. The states of these channels evolve as independent and identically distributed Markov processes (the Gilbert-Elliot channel model). A user, with limited sensing and access capability, chooses one channel to sense and access in each slot and collects a reward determined by the state of the chosen channel. Such a problem arises in cognitive radio networks for spectrum overlay, opportunistic transmissions in fading environments, and resource-constrained jamming and anti-jamming. The objective of this report is to characterize the optimal performance of such systems. The problem can be generally formulated as obtaining the maximum expected long-term reward of a partially observable Markov decision process or a restless multi-armed bandit process, for which analytical characterizations are rare. Exploiting the structure and optimality of the myopic channel selection policy established recently we obtain a closed-form expression of the maximum link throughput for two-channel systems and lower and upper bounds when there are more than two channels. These results allow us to study the rate at which the optimal performance of an opportunistic system increases with the number of channels and to obtain the limiting performance as the number of channels approaches to infinity.

15. SUBJECT TERMS

Opportunistic access, cognitive radio, spectrum overlay, dynamic channel selection, myopic policy

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**23**	

I. INTRODUCTION

The fundamental idea of opportunistic communications is to adapt the transmission parameters (data rate, modulation, transmission power, etc) according to the state of the communication environment including, for example, fading conditions, interference level, and buffer state. Since the seminal work by Knopp and Humblet in 1995 [1], the concept of opportunistic communications has found applications beyond transmission over fading channels. An emerging application is cognitive radios for spectrum overlay (also referred to as opportunistic spectrum access), where secondary users search in the spectrum for idle channels temporarily unused by primary users [2]. Another application is resource-constrained jamming and anti-jamming, where a jammer seeks channels occupied by users or a user tries to avoid jammers.

We take a simplified model of these opportunistic communication systems with N parallel channels. These N channels are modeled as independent and identically distributed Gilbert-Elliot channels [3] as illustrated in Fig. 1. The state of a channel — "good" (1) or "bad" (0) — indicates the desirability of accessing this channel and determines the resulting reward. With limited sensing and access capability, a user chooses one of the channels to sense and access in each slot, aiming to maximize its expected long-term reward (*i.e.,* , throughput). The objective of this report is to characterize analytically the maximum throughput of such a system. In particular, we are interested in the relationship between the maximum throughput and the number of channels.
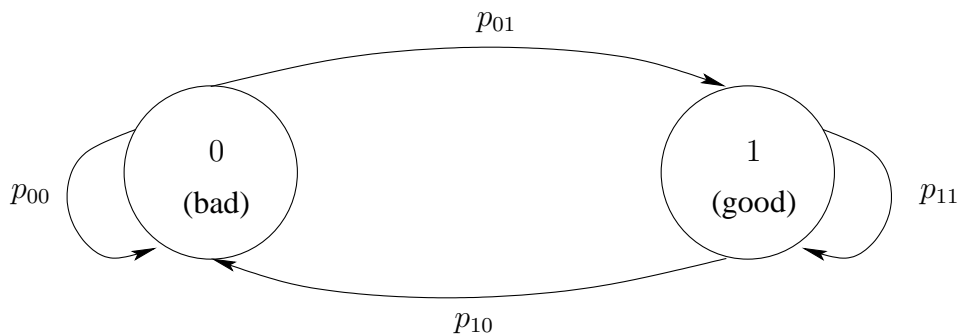

Fig. 1. The Gilbert-Elliot channel model.

This problem can be treated as a partially observable Markov decision process (POMDP) [4] or more specifically, a restless multi-armed bandit process [5] due to the independence across

channels. The maximum throughput of the multi-channel opportunistic system is essentially the maximum expected total reward, or the value function, of a POMDP [6]. Unfortunately, obtaining optimal solutions to POMDPs, even numerically, is often intractable, and closed-form expressions for value functions are rare.

In this report, we obtain a closed-form expression of the maximum throughput for two-channel opportunistic systems. For systems with more than two channels, we develop lower and upper bounds that monotonically tighten as the number N of channels increases. These results allow us to study the rate at which the optimal performance of an opportunistic system increases with N and to obtain the limiting performance as N approaches to infinity. They demonstrate that the optimal link throughput of a multi-channel opportunistic system with limited sensing quickly saturates as the number of channel increases.

Our analysis hinges on the structure and optimality of the myopic policy established in [7], [8]. The optimality of the myopic policy makes it sufficient to obtain the maximum throughput from the performance of the myopic policy, and the simple structure of the myopic policy makes it possible to characterize analytically its performance. Specifically, based on the structure of the myopic policy, we show that the performance of the myopic policy is determined by the steady-state distributions of a discrete random process with countable sample space. For $N = 2$, this random process is a first-order Markov chain. We obtain the stationary distribution of this Markov chain in closed-form, leading to exact characterizations of the maximum throughput. For $N > 2$, we construct first-order Markov processes that stochastically dominate or are dominated by the discrete random process. The stationary distributions of the former, again obtained in closed-forms, lead to lower and upper bounds on the maximum throughput.

II. PROBLEM FORMULATION

We consider the scenario where a user is trying to access the wireless spectrum using a slotted transmission structure. The spectrum consists of N independent and statistically identical channels. The state $S_i(t)$ of channel i in slot t is given by a two-state discrete-time Markov chain shown in Fig. 1.

At the beginning of each slot, the user selects one of the N channels to sense. If the channel is sensed to be in the "good" state (state 1), the user transmits and collects one unit of reward. Otherwise the user does not transmit (or transmits at a lower rate), collects no reward, and waits

until the next slot to make another choice. The objective is to maximize the average reward (throughput) over a horizon of T slots by choosing judiciously a sensing policy that governs channel selection in each slot.

Due to limited sensing, the system state $[S_1(t), \cdots, S_N(t)] \in \{0, 1\}^N$ in slot t is not fully observable to the user. It can, however, infer the state from its decision and observation history. It has been shown that a sufficient statistic of the system for optimal decision making is given by the conditional probability that each channel is in state 1 given all past decisions and observations [4]. Referred to as the belief vector, this sufficient statistic is denoted by $\Omega(t) \triangleq [\omega_1(t), \cdots, \omega_N(t)]$, where $\omega_i(t)$ is the conditional probability that $S_i(t) = 1$. Given the sensing action a and the observation S_a in slot t, the belief vector for slot $t + 1$ can be obtained as follows.

$$
\omega_i(t+1) = \begin{cases} p_{11}, & a = i, S_a = 1 \\ p_{01}, & a = i, S_a = 0 \\ \omega_i(t)p_{11} + (1 - \omega_i(t))p_{01}, & a \neq i \end{cases} . \tag{1}
$$

A sensing policy π specifies a sequence of functions $\pi = [\pi_1, \pi_2, \cdots, \pi_T]$ where π_t maps a belief vector $\Omega(t)$ to a sensing action $a(t) \in \{1, \cdots, N\}$ for slot t. Multi-channel opportunistic access can thus be formulated as the following stochastic control problem.

$$
\pi^* = \arg \max_{\pi} \mathbb{E}_{\pi} \left[\sum_{t=1}^{T} R(\pi_t(\Omega(t))) | \Omega(1) \right],
$$

where $R(\pi_t(\Omega(t)))$ is the reward obtained when the belief is $\Omega(t)$ and channel $\pi_t(\Omega(t))$ is selected, and $\Omega(1)$ is the initial belief vector. If no information on the initial system state is available, each entry of $\Omega(1)$ can be set to the stationary distribution ω_o of the underlying Markov chain:

$$
\omega_o = \frac{p_{01}}{p_{01} + p_{10}}. \tag{2}
$$

III. STRUCTURE AND OPTIMALITY OF MYOPIC POLICY

A. The Value Function

Let $V_t(\Omega)$ be the value function, which represents the maximum expected total reward that can be obtained starting from slot t given the current belief vector Ω. Given that the user takes action a and observes S_a, the reward that can be accumulated starting from slot t consists of two parts: the immediate reward $R_a(\Omega) = \omega_a$ and the maximum expected future reward $V_{t+1}(\mathcal{T}(\Omega|a, s_a))$,

where $\mathcal{T}(\Omega|a, s_a)$ denotes the updated belief vector for slot $t + 1$ as given in (1). Averaging over all possible observations S_a and maximizing over all actions a, we arrive at the following optimality equation.

$$
\begin{aligned}
V_T(\Omega) &= \max_{a=1,\cdots,N} \omega_a \\
V_t(\Omega) &= \max_{a=1,\cdots,N} (\omega_a + \omega_a V_{t+1}\left(\mathcal{T}\left(\Omega|a, 1\right)\right)) + (1 - \omega_a)V_{t+1}\left(\mathcal{T}\left(\Omega|a, 0\right)\right).
\end{aligned} \tag{3}
$$

In theory, the optimal policy π^* and its performance $V_1(\Omega_o)$ can be obtained by solving the above dynamic programming. Unfortunately, due to the impact of the current action on the future reward and the uncountable space of the belief vector Ω, obtaining the optimal solution using directly the above recursive equations is computationally prohibitive. Even when approximate numerical solutions can be obtained, they do not provide insight into system design or analytical characterizations of the optimal performance $V_1(\Omega(1))$.

B. The Myopic Policy

A myopic policy ignores the impact of the current action on the future reward, focusing solely on maximizing the expected immediate reward $R(\Omega)$. Myopic policies are thus stationary. The myopic action $\hat{a}$ under belief state $\Omega = [\omega_1, \cdots, \omega_N]$ is simply given by

$$
\hat{a}(\Omega) = \arg \max_{a=1,\cdots,N} \omega_a. \tag{4}
$$

In general, obtaining the myopic action in each slot requires the recursive update of the belief vector Ω as given in (1), which requires the knowledge of the transition probabilities $\{p_{ij}\}$. Interestingly, it has been shown in [7], [9] that the myopic policy has a simple structure that does not need the update of the belief vector or the precise knowledge of the transition probabilities.

The basic structure of the myopic policy is a round-robin scheme based on a circular ordering of the channels. For $p_{11} \geq p_{01}$, the circular order is constant and determined by a descending order of the initial belief values. The myopic action is to stay in the same channel when it is good (state 1) and switch to the next channel in the circular order when it is bad. In the case of $p_{11} < p_{01}$, the circular order is reversed in every slot with the initial order determined by the initial belief values. The myopic policy stays in the same channel when it is bad; otherwise, it switches to the next channel in the current circular order.

Another way to see the channel switching structure of the myopic policy is through the last visit to each channel (once every channel has been visited at least once). Specifically, for $p_{11} \geq p_{01}$, when a channel switch is needed, the policy selects the channel visited the longest time ago. For $p_{11} < p_{01}$, when a channel switch is needed, the policy selects, among those channels to which the last visit occurred an even number of slots ago, the one most recently visited. If there are no such channels, the user chooses the channel visited the longest time ago.

Note that the above simple structure of the myopic policy reveals that other than the order of p_{11} and p_{01}, the knowledge of the transition probabilities are unnecessary.

Surprisingly, the myopic policy with such a simple and robust structure achieves the optimal performance for $N = 2$ [7], [9]. It has been conjectured in [7], [9] (based on numerical examples[1]) that the optimality of the myopic policy can be generalized to $N > 2$. In a recent work [8], the optimality of the myopic policy has been established for a general N under the condition of $p_{11} > p_{01}$.

C. Simulation Examples

1) Figure 2 below shows the throughput (average reward per slot) as a function of time, where $N = 10$, $p_{11} = 0.1, p_{01} = 0.9$. The throughput achieved by the myopic policy increases with time, which results from the improved information on the channel state drawn from accumulating observations. This demonstrates that the myopic policy can learn from observations and track channels with the good state more effectively as the observations accumulate. Up to 50% gain can be achieved over random sensing whose performance is static with time.

2) Another example is shown in Figure 3, where we assume the channel transition probabilities change from $p_{01} = 0.1, p_{11} = 0.6$ to $p_{01} = 0.4, p_{11} = 0.9$ at $t = 6$. Note that after the change, each channel is more likely to be in the good state. From Figure 3, we can see that the myopic policy can track this change in the system model; the throughput improves significantly after $t = 5$.

---

[1]Among extensive examples, p_{01} and p_{11} are randomly chosen from interval $[0, 1]$, N is chosen between 3 and 7, and T is chosen between 1 and 20. We compare the myopic actions with the optimal actions in each example, which shows that the myopic policy is optimal.
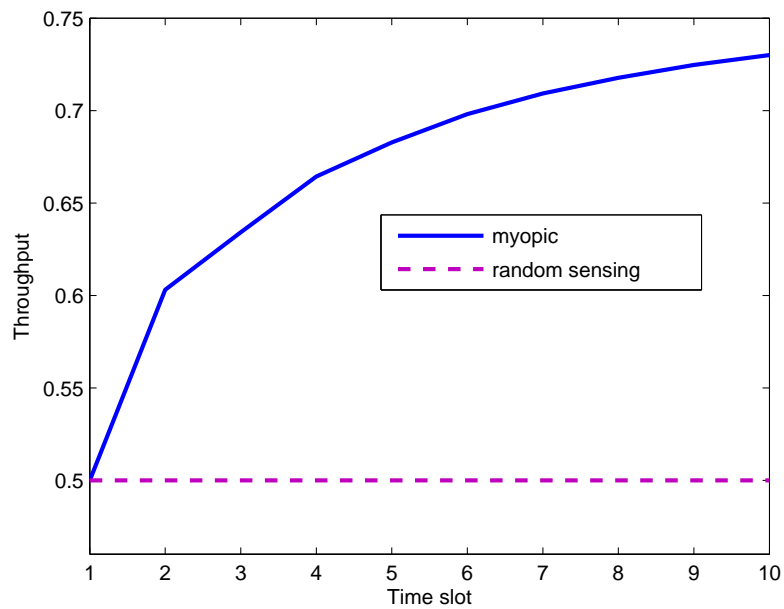
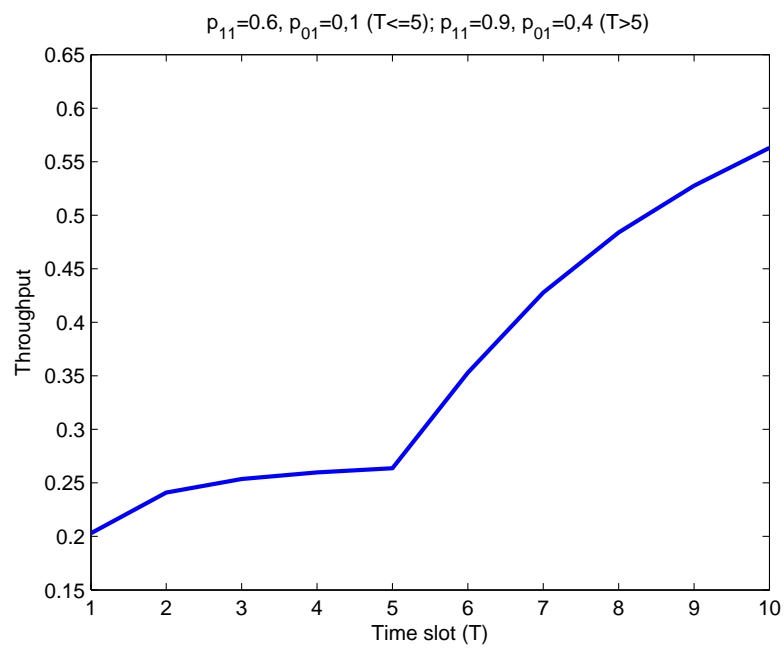Fig. 2. myopic policy v.s. random sensing policy.

Fig. 3. Tracking the change in channel transition probabilities occurred at $t = 6$.

IV. LINK THROUGHPUT LIMITS

The objective here is to characterize the link throughput limit U of multi-channel opportunistic access with limited sensing.

A. *Uniqueness of Steady-State Performance and Its Numerical Evaluation*

We first establish the existence and uniqueness of the system steady states under the myopic policy. The steady-state throughput of the myopic policy is given by

$$U(\Omega(1)) \triangleq \lim_{T \to \infty} \frac{\hat{V}_{1:T}(\Omega(1))}{T}, \tag{5}$$

where $\hat{V}_{1:T}(\Omega(1))$ is the expected total reward obtained in T slots under the myopic policy when the initial belief is $\Omega(1)$.

The simple structure of the myopic policy allows us to work with a Markov reward process with a finite state space instead of one with an uncountable state space (*i.e.,* belief vectors) as we encounter in a general POMDP. Details are stated in the Theorem below.

Theorem 1: Let $S^{(i)}(t)$ denote the state of the i-th channel in the current circular order $\mathcal{K}(t)$, where the starting point of the circular order is fixed to the myopic action, *i.e.,* $\hat{a}(t) = 1$ for all t. Then $\vec{\mathbf{S}}(t) \triangleq [S^{(1)}(t), S^{(2)}(t), \cdots, S^{(N)}(t)]\}$ forms a 2^N-state Markov chain with transition probabilities $\{q_{\vec{i},\vec{j}}\}$ given in (6), and the performance of the myopic policy is determined by the Markov reward process $(\vec{\mathbf{S}}(t), R(t))$ with $R(t) = S^{(1)}(t)$.

$$p_{11} \geq p_{01} \qquad\qquad\qquad\qquad p_{11} < p_{01}$$

$$q_{\vec{i},\vec{j}} = \begin{cases} \prod_{k=1}^{N} p_{i_k,j_k} & \text{if } i_1 = 1 \\ p_{i_1,j_N} \prod_{k=2}^{N} p_{i_k,j_{k-1}} & \text{if } i_1 = 0 \end{cases} , \qquad q_{\vec{i},\vec{j}} = \begin{cases} \prod_{k=1}^{N} p_{i_k,j_{N-k+1}} & \text{if } i_1 = 1 \\ p_{i_1,j_1} \prod_{k=2}^{N} p_{i_k,j_{N-k+2}} & \text{if } i_1 = 0 \end{cases} ,\tag{6}$$

where $\vec{i} = [i_1, i_2, \cdots, i_N]$, $\vec{j} = [j_1, j_2, \cdots, j_N]$.

Proof: The proof follows directly from the structure of the myopic policy by noticing that $S^{(1)}(t)$ determines the channel ordering in $\vec{\mathbf{S}}(t+1)$ and each channel evolves as Markov chains. Specifically, for $p_{11} \geq p_{01}$, if $S^{(1)}(t) = 1$, the channel ordering in $\vec{\mathbf{S}}(t+1)$ is the same as that in $\vec{\mathbf{S}}(t)$; if $S^{(1)}(t) = 0$, the first channel in $\vec{\mathbf{S}}(t)$ is moved to the last one in $\vec{\mathbf{S}}(t+1)$ with the ordering of the rest $N-1$ channel intact. For $p_{11} < p_{01}$, if $S^{(1)}(t) = 0$, the first channel in $\vec{\mathbf{S}}(t)$ remains the first in $\vec{\mathbf{S}}(t+1)$ while the ordering of the rest channels is reversed; if $S^{(1)}(t) = 1$,

the ordering of all N channels are reversed. The transition probabilities given in (6) thus follow.

□□□

From Theorem 1, $U(\Omega(1))$ is determined by the Markov reward process $\{\vec{\mathbf{S}}(t), R(t)\}$. It is easy to see that the 2^N-state Markov chain $\{\vec{\mathbf{S}}(t)\}$ is irreducible and aperiodic, thus has a limiting distribution. As a consequence, the limit in (5) exists, and the steady-state throughput U is independent of the initial belief value $\Omega(1)$.

Theorem 1 also provides a numerical approach to evaluating U by calculating the limiting (stationary) distribution of $\{\vec{\mathbf{S}}(t)\}$ whose transition probabilities are given in (6). Specifically, the throughput U is given by the summation of the limiting probabilities of those 2^{N-1} states with first entry $S^{(1)} = 1$. This numerical approach, however, does not provide an analytical characterization of the throughput U in terms of the number N of channels and the transition probabilities $\{p_{i,j}\}$. In the next section, we obtain analytical expressions of U and its scaling behavior with respect to N based on a stochastic dominance argument.

B. Analytical Characterization of Throughput

Our analysis hinges on the structure and optimality of the myopic policy given in Sec. III-B. The optimality of the myopic policy makes it sufficient to obtain U from the performance of the myopic policy, and the simple structure of the myopic policy makes it possible to characterize analytically its performance.

1) Transmission Period: From the structure of the myopic policy we can see that the key to the throughput is how often the user switches channels, or equivalently, how long it stays in the same channel. When $p_{11} \geq p_{01}$, the event of channel switch is equivalent to a slot *without* reward. The opposite holds when $p_{11} < p_{01}$: a channel switch corresponds to a slot *with* reward.
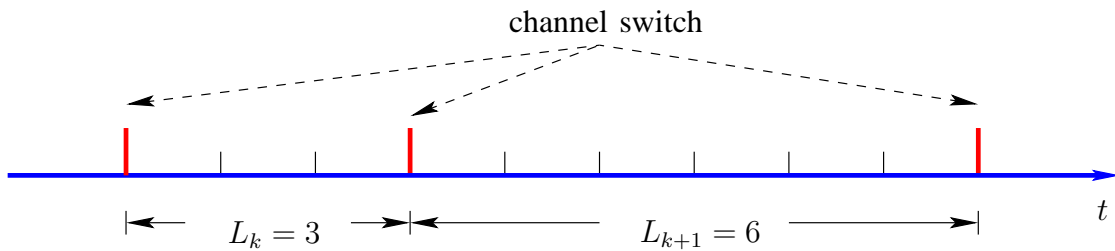

Fig. 4.   The transmission period structure.

We thus introduce the concept of transmission period, which is the time the user stays in the same channel, as illustrated in Fig. 4. Let L_k denote the length of the kth transmission period. We thus have a discrete-time random process $\{L_k\}_{k=1}^{\infty}$ with a sample space of positive integers. It is easy to show that throughput U is determined by the average length $\bar{L}$ of a transmission period as given in Lemma 1 below.

Lemma 1: Let $\bar{L} = \lim_{k\to\infty} \frac{\Sigma_{k=1}^{K} L_k}{K}$ denote the average length of a transmission period. The throughput limit U is given by

$$U = \begin{cases} 1 - 1/\bar{L}, & p_{11} \geq p_{01} \\ 1/\bar{L}, & p_{11} < p_{01} \end{cases}. \tag{7}$$

Proof: When $p_{11} \geq p_{01}$, the user collects $(L_k - 1)$ units of reward during each transmission period L_k, obtain U as the average reward over an infinite number of transmission periods. We have

$$U = \lim_{K\to\infty} \frac{\Sigma_{k=1}^{K}(L_k - 1)}{\Sigma_{k=1}^{K} L_k} = 1 - \frac{1}{\lim_{K\to\infty} \frac{\Sigma_{k=1}^{K} L_k}{K}} = 1 - \frac{1}{\bar{L}}, \tag{8}$$

where $\bar{L}$ denotes the average length of a transmission period.

When $p_{11} < p_{01}$, the user collects 1 unit of reward during each transmission period.

$$U = \lim_{K\to\infty} \frac{\Sigma_{k=1}^{K} 1}{\Sigma_{k=1}^{K} L_k} = \frac{1}{\lim_{K\to\infty} \frac{\Sigma_{i=1}^{K} L_k}{K}} = \frac{1}{\bar{L}}. \tag{9}$$

□□□

C. Link Throughput Limit for $N = 2$

For $N = 2$, $\{L_k\}_{k=1}^{\infty}$ is a first-order Markov chain. We have the following lemma.

Lemma 2: $\{L_k\}_{k=1}^{\infty}$ is an irreducible, recurrent, and aperiodic first-order Markov chain with the following unique stationary distribution (the limiting distribution) $\{\lambda_l\}_{l=1}^{\infty}$.

- Case 1: $p_{11} \geq p_{01}$

$$\lambda_l = \begin{cases} 1 - \bar{\omega}, & l = 1 \\ \bar{\omega} p_{11}^{l-2} p_{10}, & l \geq 2 \end{cases}, \tag{10}$$

where $\bar{\omega}$ is the expected probability that the channel we switch to is in state 1, *i.e.,* , the expected belief value of the channel we switch to. It is given by

$$\bar{\omega} = \frac{p_{01}^{(2)}}{1 + p_{01}^{(2)} - A}, \tag{11}$$

where $p_{01}^{(2)} = p_{00}p_{01} + p_{01}p_{11}, \; A = \frac{p_{01}}{1+p_{01}-p_{11}}(1 - \frac{(p_{11}-p_{01})^3(1-p_{11})}{1-(p_{11})^2+p_{11}p_{01}})$.

- Case 2: $p_{11} < p_{01}$

$$\lambda_l = \begin{cases} \bar{\omega}', & l = 1 \\ (1 - \bar{\omega}')p_{00}^{l-2}p_{01}, & l \geq 2 \end{cases}, \tag{12}$$

where $\bar{\omega}'$ is the expected probability that the channel we switch to is in state 1. It is given by

$$\bar{\omega}' = \frac{B}{1 - p_{11}^{(2)} + B}, \tag{13}$$

where $p_{11}^{(2)} = p_{10}p_{01} + p_{11}p_{11}, \; B = \frac{p_{01}}{1+p_{01}-p_{11}}(1 + \frac{(p_{11}-p_{01})^3(1-p_{11})}{1-(1-p_{01})(p_{11}-p_{01})})$.

Proof: Since $\{L_k\}_{k=1}^{\infty}$ is an irreducible, recurrent, and aperiodic first-order Markov Chain, if there exists a stationary distribution $\overrightarrow{\lambda} = [\lambda_1, .., \lambda_i, ...]$, then $\overrightarrow{\lambda}$ is the limiting distribution.

Case 1: $p_{11} \geq p_{01}$

The transition matrix $Q = \{q_{ij}\}$ of the Markov chain $\{L_k\}_{k=1}^{\infty}$ is

$$\begin{cases} q_{i1} = 1 - p_{01}^{(i+1)}, & i \geq 1 \\ q_{ij} = p_{01}^{(i+1)}p_{11}^{j-2}p_{10}, & i \geq 1, j \geq 2. \end{cases} \tag{14}$$

Let $Q(:, k)$ denote the kth column of Q. We have

$$\mathbf{1} - Q(:, 1) = \frac{Q(:, 2)}{p_{10}}, \tag{15}$$

where $\mathbf{1}$ is the unit column vector $[1, 1, ...]^T$. Based on the definition of stationary distribution, we have

$$\overrightarrow{\lambda}Q(:, 1) = \lambda_1 \tag{16}$$

$$\overrightarrow{\lambda}Q(:, 2) = \lambda_2 \tag{17}$$

Combine (15)-(17), we have:

$$\lambda_1 = 1 - \frac{\lambda_2}{(1 - p_{11})} \tag{18}$$

For $k \geq 2$, we have $Q(:,k) = Q(:,2)(p_{11})^{k-2}$. Together with the following equations

$$\overrightarrow{\lambda} Q(:,k) = \lambda_k, \tag{19}$$

$$\overrightarrow{\lambda} Q(:,2) = \lambda_2, \tag{20}$$

we obtain

$$\lambda_k = \lambda_2 p_{11}^{k-2} \tag{21}$$

Substituting (19) and (21) into (20), we have $[1 - \frac{\lambda_2}{1-p_{11}}, \lambda_2, \lambda_2 p_{11}, \lambda_2 p_{11}^2, \cdots] Q(:,2) = \lambda_2$. Solving for λ_2, we have $\lambda_2 = \bar{\omega} p_{10}$, which gives us the stationary distribution as

$$\lambda_k = \begin{cases} 1 - \bar{\omega}, & k = 1 \\ \bar{\omega} p_{11}^{k-2} p_{10}, & k > 1, \end{cases} \tag{22}$$

where $\bar{\omega} = \frac{p_{01}^{(2)}}{1 + p_{01}^{(2)} - A}$, and $A = \frac{p_{01}}{1 + p_{01} - p_{11}}(1 - \frac{(p_{11} - p_{01})^3(1 - p_{11})}{1 - (p_{11})^2 + p_{11} p_{01}})$.

Case 2: $p_{11} < p_{01}$

The transition matrix $Q = \{q_{ij}\}$ of the Markov chain $\{L_k\}_{k=1}^{\infty}$ is

$$\begin{cases} q_{i1} = p_{11}^{(i+1)}, & i \geq 1 \\ q_{ij} = p_{10}^{(i+1)}(p_{00})^{j-2} p_{01}, & i \geq 1, j \geq 2 \end{cases} \tag{23}$$

Similar to Case 1, we can obtain the stationary distribution $\overrightarrow{\lambda}$ of Q as

$$\lambda_k = \begin{cases} \bar{\omega}', & k = 1 \\ (1 - \bar{\omega}') p_{00}^{k-2} p_{01}, & k > 1, \end{cases} \tag{24}$$

where $\bar{\omega}' = \frac{B}{1 - p_{11}^{(2)} + B}$, and $B = \frac{p_{01}}{1 + p_{01} - p_{11}}(1 + \frac{(p_{11} - p_{01})^3(1 - p_{11})}{1 - (1 - p_{01})(p_{11} - p_{01})})$.

▢▢▢

From Lemma 2 and Lemma 1, we obtain the throughput limit U for $N = 2$ as given in the theorem below.

Theorem 2: For $N = 2$, the throughput limit U is given by

$$U = \begin{cases} 1 - \frac{1-p_{11}}{1+\bar{\omega}-p_{11}}, & p_{11} \geq p_{01} \\ \frac{p_{01}}{1-\bar{\omega}'+p_{01}}, & p_{11} < p_{01} \end{cases}, \tag{25}$$

where $\bar{\omega}$ and $\bar{\omega}'$ are given, respectively, in (22) and (24).

D. Link Throughput Limit for $N > 2$

For $N > 2$, it is difficult to obtain the average length $\bar{L}$ of a transmission period. Our objective is to develop lower and upper bounds on the throughput limit U.

The approach is to construct first-order Markov chains that stochastically dominate or are dominated by $\{L_k\}_{k=1}^{\infty}$. The limiting distributions of these first-order Markov chains, which can be obtained in closed-form, thus lead to lower and upper bounds on U according to Lemma 1. Specifically, for $p_{11} \geq p_{01}$, a lower bound on U is obtained by constructing a first-order Markov chain whose limiting distribution is stochastically dominated by the stationary distribution of $\{L_k\}_{k=1}^{\infty}$. An upper bound on U is given by a first-order Markov chain whose stationary distribution stochastically dominates the stationary distribution of $\{L_k\}_{k=1}^{\infty}$. Similarly, bounds on U for $p_{11} < p_{01}$ can be obtained.

Theorem 3: For $N > 2$, we have the following lower and upper bounds on the throughput limit U.

- Case 1: $p_{11} \geq p_{01}$

$$\frac{C}{C + (1 - D + C)(1 - p_{11})} \leq U \leq \frac{\omega_o}{1 - p_{11} + \omega_o}, \tag{26}$$

where ω_o is given by (2), $C = \omega_o(1 - (p_{11} - p_{01})^N)$, $D = \omega_o(1 - \frac{(p_{11}-p_{01})^{N+1}(1-p_{11})}{1-(p_{11})^2+p_{11}p_{01}})$.

- Case 2: $p_{11} < p_{01}$

$$1 - \frac{p_{10}^{(2)}}{E - p_{01}H} \leq U \leq 1 - \frac{p_{10}^{(2)}}{E - p_{01}G} \tag{27}$$

where $p_{10}^{(2)} = p_{10}p_{00} + p_{11}p_{10}$,

$$E = p_{10}^{(2)}(1 + p_{01}) + p_{01}(1 - F),$$

$$F = (1 - p_{01})(1 - \omega_o)\left(\frac{1}{2-p_{01}} - \frac{p_{01}(p_{11}-p_{01})^4}{1-(p_{11}-p_{01})^2(1-p_{01})^2}\right),$$

$$G = (1 - \omega_o)\left(\frac{1}{2-p_{01}} - \frac{p_{01}(p_{11}-p_{01})^6}{1-(p_{11}-p_{01})^2(1-p_{01})^2}\right),$$

$$H = (1 - \omega_o)\left(\frac{1}{2-p_{01}} - \frac{p_{01}(p_{11}-p_{01})^{2N-1}}{1-(p_{11}-p_{01})^2(1-p_{01})^2}\right).$$

- Monotonicity: for both cases, the difference between the upper and lower bounds monotonically decreases with N; for $p_{11} \geq p_{01}$, the lower bound converges to the upper bound as $N \to \infty$.

Proof:

Case 1: $p_{11} \geq p_{01}$

Let ω_k denote the belief value of the chosen channel in the first slot of the k-th TP. The length $L_k(\omega_k)$ of this TP has the following distribution.

$$\Pr[L_k(\omega_k) = l] = \begin{cases} 1 - \omega_k, & l = 1 \\ \omega_k p_{11}^{l-2}p_{10}, & l > 1 \end{cases}. \tag{28}$$

It is easy to see that if $\omega' \geq \omega$, then $L_k(\omega')$ stochastically dominates $L_k(\omega)$.
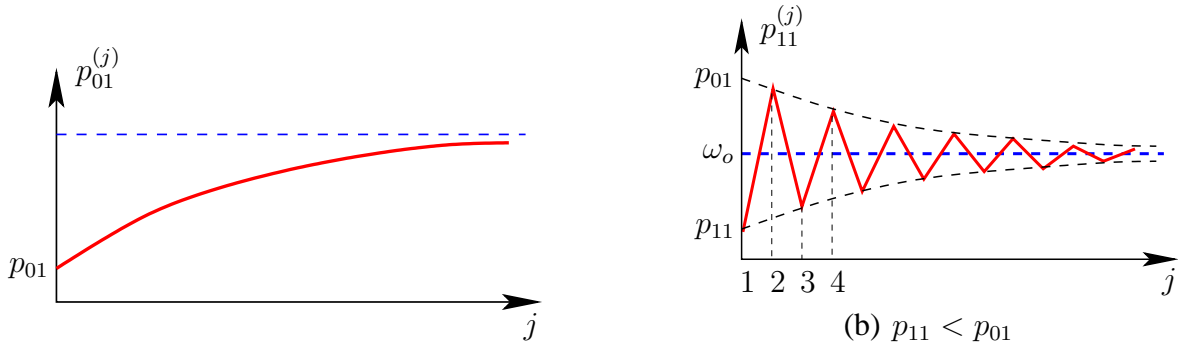

Fig. 5. The j-step transition probabilities of the Gilbert-Elliot channel.

From the round-robin structure of the myopic policy, $\omega_k = p_{01}^{(J_k)}$, where $J_k = \sum_{i=1}^{N-1} L_{k-i} + 1$. Based on the monotonic increasing property of the j-step transition probability $p_{01}^{(j)}$ (see Fig. 5), we have $\omega_k \leq \omega_o$, where ω_o is the stationary distribution of the Gilbert-Elliot channel given in

(2). $L_k(\omega_o)$ thus stochastically dominates $L_k(\omega_k)$, and the expectation of the former, $\overline{L_k(\omega_o)} = 1 + \frac{\omega_o}{1-p_{11}}$, leads to the upper bound of U given in (26).

Next, we prove the lower bound of U by constructing a hypothetical system where the initial belief value of the chosen channel in a TP is a lower bound of that in the real system. The average TP length in this hypothetical system is thus smaller than that in the real system, leading to a lower bound on U based on (7). Specifically, since $\omega_k = p_{01}^{(J_k)}$ and $J_k = \sum_{i=1}^{N-1} L_{k-i} + 1 \geq N + L_{k-1} - 1$, we have $\omega_k \leq p_{01}^{(N+L_{k-1}-1)}$. We thus construct a hypothetical system given by a first-order Markov chain $\{L_k'\}_{k=1}^{\infty}$ with the following transition probability $r_{i,j}$.

$$r_{i,j} = \begin{cases} 1 - p_{01}^{(N+i-1)}, & i \geq 1, \ j = 1 \\ p_{01}^{(N+i-1)}(p_{11})^{j-2}p_{10}, & i \geq 1, \ j \geq 2 \end{cases}. \tag{29}$$

Lemma 3: The stationary distribution of the first order Markov chain $\{L_k'\}_{k=1}^{\infty}$ is stochastically dominated by the stationary distribution of $\{L_k\}_{k=1}^{\infty}$.

Proof:

Let ω_k' denote the expected probability that the chosen channel is in state 1 in the first slot of the k-th transmission period of $\{L_k'\}_{k=1}^{\infty}$. Assume in the k-th transmission period, the distributions of L_k' and L_k both equal to the same distribution $\overrightarrow{\lambda}$, which may or may not be the statrionary distribution of $\{L_k\}_{k=1}^{\infty}$. Next we show $\omega_{k+n} \geq \omega_{k+n}'$ for any $n \geq 1$ by induction.

When $n = 1$, we have

$$\begin{aligned} \omega_{k+1} &= \Sigma_{l=1}^{\infty} \mathbb{E}_{L_{k-N+2},\cdots,L_{k-1}}[p_{01}^{1+\Sigma_{i=k-N+2}^{k}L_i}|L_k = l]Pr(L_k = l) \\ &\geq \Sigma_{l=1}^{\infty} \mathbb{E}_{L_{k-N+2},\cdots,L_{k-1}}[p_{01}^{N-1+L_k}|L_k = l]Pr(L_k = l) \\ &= \Sigma_{l=1}^{\infty} p_{01}^{N-1+l}\lambda_l \\ &= \omega_{k+1}'. \end{aligned} \tag{30}$$

Assume $\omega_{k+n} \geq \omega_{k+n}'$, then

$$\begin{aligned} \omega_{k+n+1} &= \Sigma_{l=1}^{\infty} \mathbb{E}_{L_{k+n-N+2},\cdots,L_{k+n-1}}[p_{01}^{1+\Sigma_{i=k+n-N+2}^{k+n}L_i}|L_{k+n} = l]Pr(L_{k+n} = l) \\ &\geq \Sigma_{l=1}^{\infty} \mathbb{E}_{L_{k+n-N+2},\cdots,L_{k+n-1}}[p_{01}^{N-1+L_{k+n}}|L_{k+n} = l]Pr(L_{k+n} = l) \\ &= \Sigma_{l=1}^{\infty} p_{01}^{N-1+l}Pr(L_{k+n} = l) \end{aligned} \tag{31}$$

Since $\omega_{k+n} \geq \omega'_{k+n}$, by (28), we have

$$Pr(L_{k+n} = l) \leq Pr(L'_{k+n} = l), \quad if \quad l = 1;$$

$$Pr(L_{k+n} = l) \geq Pr(L'_{k+n} = l), \quad if \quad l > 1. \tag{32}$$

Since the largest number in the series $\{p_{01}^{N-1+l}\}_{l=1}^{\infty}$ is the first one, by (32) and the fact that $\Sigma_{l=1}^{\infty} Pr(L_{k+n} = l) = \Sigma_{l=1}^{\infty} Pr(L'_{k+n} = l) = 1$, we have

$$\Sigma_{l=1}^{\infty} p_{01}^{N-1+l} Pr(L_{k+n} = l) \geq \Sigma_{l=1}^{\infty} p_{01}^{N-1+l} Pr(L'_{k+n} = l) = \omega'_{k+n+1} \tag{33}$$

Combine (31) and (33), we have $\omega_{k+n+1} \geq \omega'_{k+n+1}$.

By the above induction, we have $\omega_{k+n} \geq \omega'_{k+n}$ for any $n \geq 1$. So the stationary distribution of the first order Markov chain $\{L'_k\}_{k=1}^{\infty}$ is dominated by the stationary distribution of $\{L_k\}_{k=1}^{\infty}$.

The first order Markov chain $\{L'_k\}_{k=1}^{\infty}$ has the following transition matrix $S = \{s_{ij}\}_{i,j=1}^{\infty}$

$$\begin{cases} s_{i1} = 1 - p_{01}^{(N+i-1)}, & i \geq 1 \\ s_{ij} = p_{01}^{(N+i-1)}(p_{11})^{j-2} p_{10}, & i \geq 1, j \geq 2 \end{cases}. \tag{34}$$

Let $\overline{L'}$ denote the average length of a transmission period of L'_k. Solving for the stationary distribution of $\{L'_k\}_{k=1}^{\infty}$ from S, we obtain $\overline{L'}$, which leads to a lower bound on U according to Lemma 3 and Lemma 1.

Case 2: $p_{11} < p_{01}$

In this case, the larger the initial belief of the chosen channel in a given TP, the smaller the average length of the TP. On the other hand, (7) shows that U is inversely proportional to the average TP length. Thus, similar to the case of $p_{11} \geq p_{01}$, we will construct hypothetical systems where the initial belief of the chosen channel in a TP is an upper bound or a lower bound of that in the real system. The former leads to an upper bound on U, the latter, a lower bound on U.

Consider first the upper bound. From the structure of the myopic policy, it is clear that when L_{k-1} is odd, in the k-th TP, the user will switch to the channel visited in the $(k-2)$-th TP. As a consequence, the initial belief ω_k of the k-th TP is given by $\omega_k = p_{11}^{(L_{k-1}+1)}$. When L_{k-1} is

even, we can show that $\omega_k \leq p_{11}^{(L_{k-1}+4)}$. This is because that for $N \geq 3$ and L_{k-1} even, the user cannot switch to a channel visited $L_{k-1}+2$ slots ago, and $p_{11}^{(j)}$ decreases with j for even j's and $p_{11}^{(j)} > p_{11}^{(i)}$ for any even j and odd i (see Fig. 5). We thus construct a hypothetical system given by the first-order Markov chain $\{L'_k\}_{k=1}^{\infty}$ with the following transition probabilities.

$$r_{i,j} = \begin{cases} p_{11}^{(i+1)}, & \text{if } i \text{ is odd, } j = 1 \\ p_{10}^{(i+1)}(p_{00})^{j-2}p_{01}, & \text{if } i \text{ is odd, } j \geq 2 \\ p_{11}^{(i+4)}, & \text{if } i \text{ is even, } j = 1 \\ p_{10}^{(i+4)}(p_{00})^{j-2}p_{01}, & \text{if } i \text{ is even, } j \geq 2 \end{cases}. \tag{35}$$

Similar to the proof of Lemma 3, it can be shown that the stationary distribution of $\{L'_k\}_{k=1}^{\infty}$ is stochastically dominated by that of $\{L_k\}_{k=1}^{\infty}$. The former leads to the upper bound of U given in (27).

We now consider the lower bound. Similarly, $\omega_k = p_{11}^{(L_{k-1}+1)}$ when L_{k-1} is odd. When L_{k-1} is even, to find a lower bound on ω_k, we need to find the smallest odd j such that the last visit to the channel chosen in the k-th TP is j slots ago. From the structure of the myopic policy, the smallest feasible odd j is $L_{k-1}+2N-3$, which corresponds to the scenario where all N channels are visited in turn from the $(k-N+1)$-th TP to the k-th TP with $L_{k-N+1} = L_{k-N+2} = \cdots = L_{k-2} = 2$. We thus have $\omega_k \geq p_{11}^{(L_{k-1}+2N-3)}$. We then construct a hypothetical system given by the first-order Markov chain $\{L'_k\}_{k=1}^{\infty}$ with the following transition probabilities.

$$r_{i,j} = \begin{cases} p_{11}^{(i+1)}, & \text{if } i \text{ is odd, } j = 1 \\ p_{10}^{(i+1)}(p_{00})^{j-2}p_{01}, & \text{if } i \text{ is odd, } j \geq 2 \\ p_{11}^{(i+2N-3)}, & \text{if } i \text{ is even, } j = 1 \\ p_{10}^{(i+2N-3)}(p_{00})^{j-2}p_{01}, & \text{if } i \text{ is even, } j \geq 2 \end{cases}. \tag{36}$$

The stationary distribution of this hypothetical system leads to the lower bound of U given in (27).

Monotonicity

The monotonicity statements can be shown by noticing that for both cases, the lower bound increases with N, while the upper bound is not a function of N.

□□□

Coroallary 1: For $p_{11} > p_{01}$, the lower bound on throughput U converges to the constant upper bound at geometrical rate $(p_{11} - p_{01})$ as N increases; for $p_{11} < p_{01}$, the lower bound on

U converges to a constant at geometrical rate $(p_{01} - p_{11})^2$.

Proof:

Let $x = |p_{11} - p_{01}|$. For $p_{11} > p_{01}$, after some simplifications, the lower bound has the form $a + b/(x^N + c)$, where a, b, c ($c \neq 0$) are constants. The upper bound is $a + b/c$. We have $\frac{|a+b/(x^N+c)-a-b/c|}{x^N} \to b/c^2$ as $N \to \infty$. Thus the lower bound converges to the upper bound with geometric rate x.

For $p_{11} < p_{01}$, the lower bound has the form $d + e/(x^{2N-1} + f)$, where d, e, f ($f \neq 0$) are constants. It converges to $d + e/f$ as $N \to \infty$. We have $\frac{|d+e/(x^{2N-1}+f)-d-e/f|}{x^{2N}} \to e/(xf^2)$ as $N \to \infty$. Thus the lower bound converges with geometric rate x^2. ☐☐☐

Though it is difficult to get a closed-form throughput limit for $N > 2$, we can calculate the throughput limit numerically by Theorem 1. We show an example of the throughput limit for $N > 2$ and $p_{11} \geq p_{01}$ as follows.
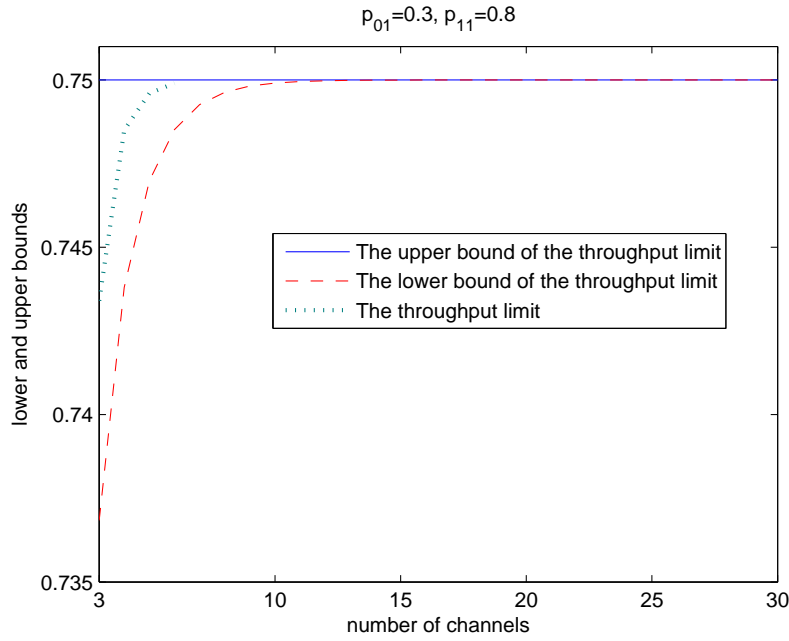

Fig. 6. The throughput limit for $N > 2$ and $p_{11} \geq p_{01}$.

E. Link Throughput over A Finite Horizon

It is interesting to note that we can obtain a closed-form expression for the throughput during a finite period T under certain conditions.

Theorem 4: When $p_{11} \geq p_{01}$ and $N \geq T$, the maximum expected total reward over T slots when the initial belief $\Omega(1)$ is given by the stationary distribution ω_o is a function of T and ω_o:

$$V(\omega_o, T) = \frac{\omega_o(T - 2)}{1 - p_{11} + \omega_o} + \frac{\omega_o(\omega_o - p_{11})^3(1 - (p_{11} - \omega_o)^{T-2})}{(1 - p_{11} + \omega_o)^2} + \omega_o + p_{11}\omega_o + (1 - \omega_o)\omega_o \quad (37)$$

Proof: From the structure of the myopic policy, if the user observes state 1 from a channel, it will stay on that channel. Otherwise, it will switch to a new channel. Clearly, V does not depend on N since at most T channels need to be <u>considered during T</u> slots.

In the first slot, the user randomly chooses one channel and gets ω_o unit of reward. Then the user will either stay or switch. This process is a Markov chain with states "stay" and "switch" as shown below.
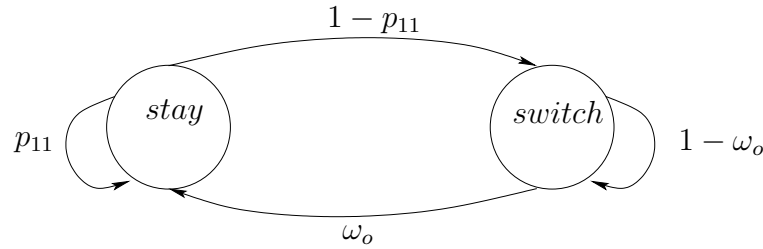

Fig. 7.   The Markov chain with states "stay" and "switch".

If the user observes 1 after the fist slot, it will stay and get p_{11} unit of reward. Otherwise it will switch to a new channel and get ω_o unit of reward. So V is determined by the distribution of the states of the above two-state Markov chain.

$$V(\omega_o, T) = \Sigma_{M=1}^{T-1} [\omega_o \quad 1 - \omega_o] \begin{bmatrix} p_{11} & 1 - p_{11} \\ \omega_o & 1 - \omega_o \end{bmatrix}^{M-1} \begin{bmatrix} p_{11} \\ \omega_o \end{bmatrix} + \omega_o$$

$$= \Sigma_{M=1}^{T-2} [\omega_o \quad 1 - \omega_o] \begin{bmatrix} p_{11} & 1 - p_{11} \\ \omega_o & 1 - \omega_o \end{bmatrix}^{M} \begin{bmatrix} p_{11} \\ \omega_o \end{bmatrix} + \omega_o + p_{11}\omega_o + (1 - \omega_o)\omega_o$$

$$= \Sigma_{M=1}^{T-2} [\omega_o \quad 1 - \omega_o] \{ \frac{1}{1 - p_{11} + \omega_o} \begin{bmatrix} \omega_o & 1 - p_{11} \\ \omega_o & 1 - p_{11} \end{bmatrix} + \frac{(p_{11} - \omega_o)^M}{1 - p_{11} + \omega_o}$$

$$\begin{bmatrix} 1 - p_{11} & p_{11} - 1 \\ -\omega_o & \omega_o \end{bmatrix} \} \begin{bmatrix} p_{11} \\ \omega_o \end{bmatrix} + \omega_o + p_{11}\omega_o + (1 - \omega_o)\omega_o$$

$$= \frac{\omega_o(T - 2)}{1 - p_{11} + \omega_o} + \frac{\omega_o(\omega_o - p_{11})^3(1 - (p_{11} - \omega_o)^{T-2})}{(1 - p_{11} + \omega_o)^2} + \omega_o + p_{11}\omega_o + (1 - \omega_o)\omega_o$$

$$\tag{38}$$

▢▢▢

From the above, we immediately see the link throughput limit U as $N \to \infty$ is given as follows:

$$U = \lim_{T \to \infty} \frac{V(\omega_o, T)}{T} = \frac{\omega_o}{1 - p_{11} + \omega_o}, \tag{39}$$

which agrees with the upper bound given in Theorem 3.

F. Numerical Examples

In this section, we demonstrate the tightness of the bounds on U given in Sec. IV-D by examining the relative difference $d(N)$ between the upper and the lower bound, where $d(N)$ is defined as the difference of the lower and upper bound divided by the upper bound. In Fig. 8, we plot $d(N = 5)$ with respect to the upper bound for $p_{11} \geq p_{01}$. From Fig. 8 we observe that for most values of p_{11} and p_{01}, $d(N = 5)$ is below 6%, demonstrating the tightness of the bounds even for a small number of channels. Furthermore, Fig. 8 shows that the bounds are tighter for larger p_{01}. Similarly observations can be drawn from Fig. 9 for the case of $p_{11} < p_{01}$.
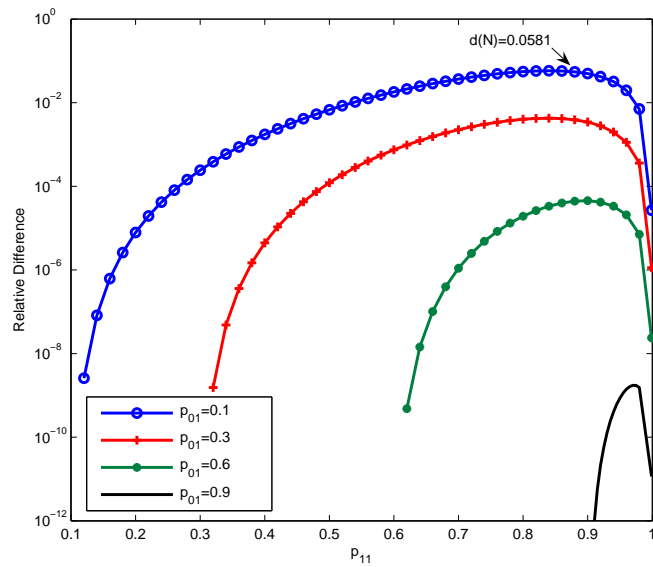
Fig. 8.   The relative difference $d(N = 5)$ between the upper and the lower bound for $p_{11} \geq p_{01}$.
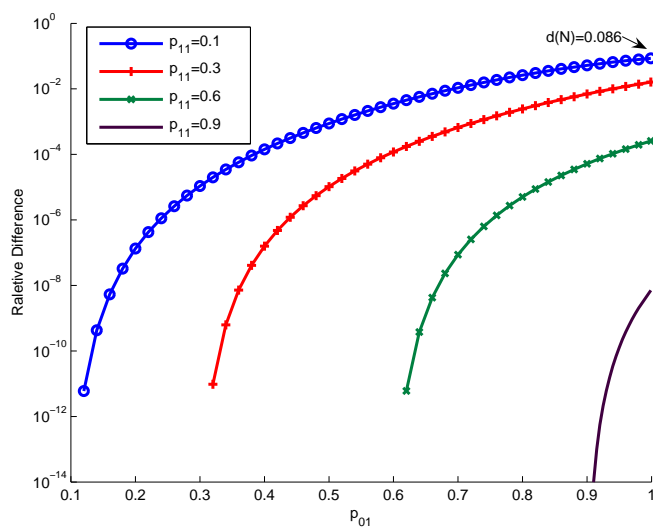

Fig. 9.   The relative difference $d(N = 5)$ between the upper and the lower bound for $p_{11} < p_{01}$.

In Fig. 10 and 11 we examine the rate at which the lower bound approaches to the upper bound as N increases. Specifically, we plot the ratio of $d(N = 10)$ to $d(N = 3)$. We observe that in both cases, the lower bound approaches to the upper bound quickly. While demonstrating the usefulness of the bounds for small N, this observation conveys a pessimistic message: the optimal link throughput of a multi-channel opportunistic system with limited sensing quickly saturates as N increases.
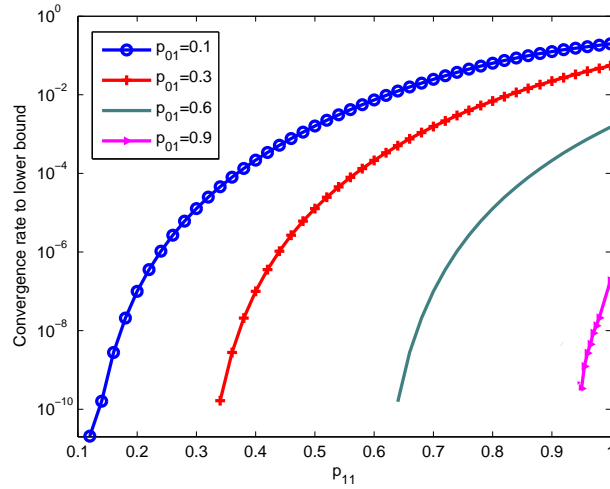

Fig. 10. The rate at which the lower bound approaches to the upper bound as N increases ($p_{11} \geq p_{01}$).
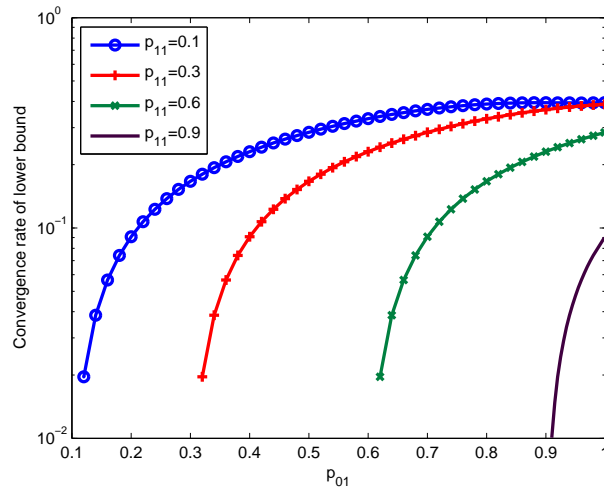

Fig. 11. The rate at which the lower bound approaches to the upper bound as N increases ($p_{11} < p_{01}$).

V. CONCLUSION AND FUTURE WORK

In this report, we have analyzed the optimal link throughput of multi-channel opportunistic communication systems under an i.i.d. Gilbert-Elliot channel model. The obtained analytical results allow us to systematically examine the impact of the number of channels and channel dynamics (transition probabilities) on the system performance. Future work includes the generalization to cases with sensing errors and non-identical channels. The former can again be addressed by exploiting the structure and optimality of the myopic policy in the presence of sensing errors as established in [10].

REFERENCES

[1] R. Knopp and P. Humblet, "Information capacity and power control in single cell multi-user communications," in *Proc. Intl Conf. Comm.*, (Seattle, WA), pp. 331–335, June 1995.

[2] Q. Zhao and B. Sadler, "A Survey of Dynamic Spectrum Access: Signal Processing, Networking, and Regulatory Policy," *IEEE Signal Processing magazine*, vol. 24, pp. 79–89, May 2007.

[3] E.N. Gilbert, "Capacity of burst-noise channels," Bell Syst. Tech. J., vol. 39, pp. 1253-1265, Sept. 1960.

[4] R. Smallwood and E. Sondik, "The optimal control of partially ovservable Markov processes over a finite horizon," *Operations Research*, pp. 1071–1088, 1971.

[5] P. Whittle, "Restless bandits: Activity allocation in a changing world", in *Journal of Applied Probability*, Volume 25, 1988.

[6] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized Cognitive MAC for Opportunistic Spectrum Access in Ad Hoc Networks: A POMDP Framework," *IEEE JSAC*, April 2007.

[7] Q. Zhao and B. Krishnamachari, "Structure and optimality of myopic sensing for opportunistic spectrum access," in *Proc. of IEEE CogNet*, June, 2007.

[8] T. Javidi, B. Krishnamachari, Q. Zhao, and M. Liu, "Optimality of Myopic Sensing in Multi-Channel Opportunistic Access," in Proc. of *ICC*, May, 2008.

[9] Q. Zhao, B. Krishnamachari, and K. Liu, "On Myopic Sensing for Multi-Channel Opportunistic Access: Structure, Optimality, and Performance," submitted to IEEE Transactions on Wireless Communications in November, 2007, revised in May, 2008.

[10] Q. Zhao, B. Krishnamachari, and K. Liu, "Low-Complexity Approaches to Spectrum Opportunity Tracking," in *Proc. of CrownCom*, August 2007.